

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-081572

(43)Date of publication of application : 28.03.1997

(51)Int.Cl.

G06F 17/28

(21)Application number : 07-234030

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 12.09.1995

(72)Inventor : YOSHIMURA YUMIKO  
KINOSHITA SATOSHI

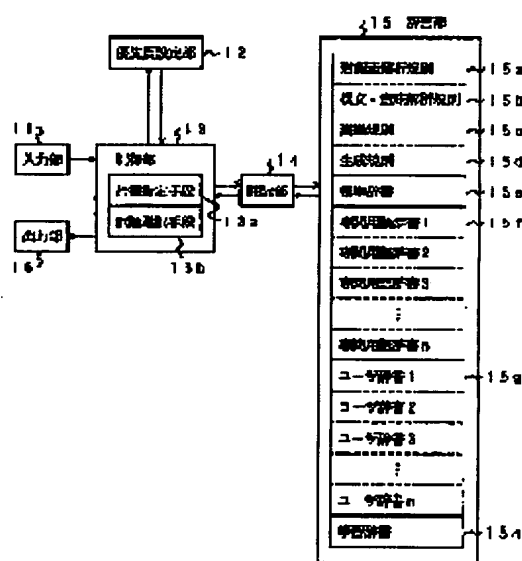
## (54) TRANSLATION DEVICE AND DICTIONARY PRIORITY SETTING METHOD

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To reduce the labor of words selection in translation and post-editing by taking the result of user's words selection in translation as learning materials to adjust the initial priority designation contents to a priority more adapted to a translation object document without user's knowing.

**SOLUTION:** A first language as the translation object is translated to a second language by a translation part 14 in accordance with classifications and priorities of dictionaries designated as optional dictionaries out of dictionaries preliminarily prepared in a dictionary part 15.

In this case, it is discriminated whether the priority change is necessary or not based on selection contents of words selection obtained as the translation result by a priority setting part 12; and if it is necessary, priorities are so adjusted that the dictionary from which words in translation are selected most frequently has the highest priority. Thus, words in translation registered in plural dictionaries are always generated in a translation with optimum priorities by user's words selection in translation, and as the result, the labor of user's words selection in translation and post-editing is reduced.



## LEGAL STATUS

[Date of request for examination] 11.09.2000

[Date of sending the examiner's decision of rejection] 04.06.2002

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]

(Translation of selected lines)

**JP H09-81572**

[0036] We will now describe a process flow in which the translation process executing unit 14 accomplishes translation. This process flow of translation is indicated in Fig.4 in form a flowchart. The translation process executing unit 14, first, obtains, by employing the control unit 13 as a mediator between the translation process executing unit 14 and the priority setting unit 12, optional class dictionary identifying codes and priority indicators indicating a priority ranking of dictionaries specified by these identifying codes both of which are held stored in the priority setting unit 12 (Step S401).

[0037] The translation process executing unit 14, after obtaining data of both these two kinds, uses specialty sector expression dictionaries 15f and user dictionaries 15g relevant to the optional class dictionary identifying codes, which is one of the two kinds of data obtained from the priority setting unit 12, together with morphologic element analysis rules 15a and common sector dictionaries 15e, which are held stored in the dictionary storage unit 15, and executes morphologic-element analysis and dictionary consultation processes (Step S402).

[0038] In the stage up to the above described steps, the translation process executing unit 14 acquires after-translation expressions from all the dictionaries as long as they contain the subject before-translation expression as their headwords respectively. These after-translation expressions collected for the same before-translation expression from various dictionaries are then subjected to a merging process, in Step S403.

[0039] In particular, the translation process executing unit 14 assembles this plurality of expressions all in correspondence to the same one headword and contained in a plurality of dictionaries respectively, and merges them as if they, as a whole, represent a set of data provided in a single dictionary for a relevant headword in a manner compliant to the priority indicators indicating priorities among the dictionaries and obtained from the priority setting unit 12.

[0040] We will now describe, as an example case, optional class dictionary identities and priority indicators of dictionaries relevant to these identities, and also an example of the result of merging

various expressions contained for a headword 'facility' within confines of thus adopted dictionaries in a manner complying to these priority indicators.

- Optional class dictionary identities

Specialty sector expression dictionaries 1, 2 and 3

User dictionaries 1 and 2

- Priority ranking of dictionaries (listed according to the associated priority ranking)

Specialty sector expression dictionaries 1, 2, 3, followed by

User dictionaries 2, 1

- Dictionary contents

Specialty sector expression dictionary 1

after-translation expressions: kinou 機能, shisetu 施設, souchi 装置

Specialty sector expression dictionary 2

after-translation expressions: kinou 機能

Specialty sector expression dictionary 3

after-translation expressions: souchi 装置, kikou 機構

User dictionary 1

after-translation expressions: fashirityi ファシリティ

User dictionary 2

after-translation expressions: setsubi 設備

Common sector dictionary

after-translation expressions:

kinou 機能, sisetsu 施設, kikou 機構, setsubi 設備, nouryoku 能力

- Result of merging process

after-translation expressions: setsubi 設備 (U2, C)

fashirityi ファシリティ (U1)

kinou 機能 (T1, T2, C)

sisetsu 施設 (T1, C)

souchi 装置 (T1, T3)

kikou 機構 (T3, C)

nouryoku 能力 (C)

Here, indicated by T1, T2 and T3 are respectively different specialty sector expression dictionaries, indicated by U1 and U2 are respectively different user dictionaries and C indicates a common sector dictionary.

[0041] In this example case described above, any after-translation expression found in a user dictionary is given a higher priority rank than any found in a specialty sector dictionary. The priority ranking of dictionaries implies a ranking among dictionaries only of the same class, i.e., one priority ranking is assigned for specialty sector dictionary class dictionaries and another priority ranking is assigned for user dictionary class dictionaries. It is obvious, however, that this configuration may be replaced by one in which a priority ranking is assigned for dictionaries of a class obtained by combining the specialty sector dictionary class and user dictionary class.

[0042] The translation process executing unit 14, according to this example case, determines at this point whether the first priority after-translation expression (setsubi 設備(U2, C)) in a result of executing the merging process matches with a learned-type after-translation expression if such a type of data as acquired through a learning process is held in a learned expression dictionary 15h according to an associated user having performed an after-translation expression selection process on an occasion preceding this point in time. And when the learned after-translation expression is determined to match with any other after-translation expression alternatives than the first priority after-translation expression, the translation process executing unit 14 comes to change the priority ranking assigned to those expressions of a set resulted from the merging process.

[0043] But if the learned after-translation expression is not found in the expressions of the set resulted from this merging process, the translation process executing unit 14 comes to execute a process for introducing this learned after-translation expression into this set of the expressions representing the result of the merging process as the first priority rate expression. If the learned expression is an expression 'kinou 機能', the expression set will be changed to the following form.

[0044] • Result of merging process

after-translation expressions:

kinou 機能 (T1, U2, C)  
setsubi 設備 (U2, C)  
fashirityi ファシリティ (U1)  
sisetsu 施設 (T1, C)  
souchi 装置 (T1, T3)  
kikou 機構 (T3, C)  
nouryoku 能力 (C)

If the learned expression is an expression ‘bengi 便宜’, the expression set will be changed to the following form.

[004 5] • Result of merging process

after-translation expressions:

bengi 便宜, setsubi 設備 (U2, C)  
fashirityi ファシリティ (U1)  
kinou 機能 (T1, U2, C)  
sisetsu 施設 (T1, C)  
souchi 装置 (T1, T3)  
kikou 機構 (T3, C)  
nouryoku 能力 (C)

Each time when completing the merging process for merging data contained in various dictionaries in the above described manner, the translation process executing unit 14 updates a frequency count table, which is a table for holding, for each of these dictionaries, a number of times dictionary headwords occurring in original sentences of a concern set (Step S404).

[0046] This frequency count table, in the case described here, is held in a work area which is, while not shown in form of a drawing, associated with the translation process executing unit 14 and this table is updated in a course of a process being executed by the translation process executing unit 14. This frequency count table, here, is initialized when newly starting a series of translation processes in Step S202 described in Fig.2.

[0047] Drawing (a) presented in Fig.7 is this frequency count table as it is at this point of the flow

of this translation work. As indicated in this table, occurrences of headwords of one dictionary in original sentences of a set of concern are counted separately between cases in which relevant after-translation expressions are rated as the first priority level and cases in which relevant after-translation expressions are rated as other than the first priority level, and a case in which the relevant after-translation expression is rated as the other than the first priority level is not included into the count if any other after-translation expression of the same headword contained in the same dictionary of concern is rated as the first priority level. Counting of the cumulative number of occurrences in this manner is continued until the translation apparatus detects an entry of an instruction for terminating the relevant translation work.

[0048] This process is followed by the translation process executing unit 14 commencing syntactic and semantic analyses according to syntactic and semantic analysis rules 15b (Step S405), and then, using thus obtained analysis results for converting the original sentences into sentences of a second language structure (Step S406). The translation process executing unit 14, then, completes the translation work of one run by generating second language sentences according to generation rules 15d (Step S407).

[0049] We will now describe a process flow, which follows an after-translation expression selection process and is concerned with the priority setting unit 12 conducting re-evaluation and revision of a priority ranking of dictionaries. Fig.5 presents a flowchart indicating a process flow associated with these processes. When a user selects an expression other than the one presented as the first priority after-translation expression, the priority setting unit 12 comes update content data of the frequency count table, which indicates the by-dictionary frequency counts, in accordance to data pieces that respectively identify the dictionary which contains thus selected after-translation expression and dictionary which contains this first priority after-translation expression (Step S501).

[0050] If these operations are accomplished, for example, in relation to the above mentioned word 'facility' and an after-translation sentence is proposed as a translation result for a subject original sentence in a form adopting an after-translation expression found in User dictionary 2 as result of it having been rated as the first priority after-translation expression, and in response to this proposed after-translation sentence the after-translation expression 'fashirity ファシリティ' has been selected

in the after-translation expression selection process, the frequency count table comes to be revised in the manner described by drawing (b) shown in Fig.7.

[0051] In this example case, the frequency count connecting to User dictionary 1 and the first priority after-translation expressions is incremented from 14 to 15 and the frequency count connecting to User dictionary 1 and other than the first priority after-translation expressions is decremented from 5 to 4, and, in contrast to these User dictionary 1 related changes, the frequency count connecting to User dictionary 2 and the first priority after-translation expression is decremented from 11 to 10 and the frequency count connecting to User dictionary 2 and other than the first priority after-translation expressions is incremented from 0 to 1.

[0052] This step of updating contents of the frequency count table is followed by the priority setting unit 12 determining whether it is necessary to amend the priority ranking of these dictionaries (Step S502).

[0053] In this particular embodiment, the priority ranking is configured so as to be amended when and if all the below listed conditions are met. Condition 1: The frequency count connecting to the dictionary which contains the after-translation expression that has been selected in the immediately preceding after-translation expression selection process and the first priority after-translation expressions is greater than the frequency count connecting to the dictionary of which the priority is ranked immediately above and the first priority after-translation expressions.

[0054] Condition 2: The number of times after-translation expressions having been selected in courses of executing the after-translation expression selection process with respect to one specific dictionary becomes greater than the frequency count connecting to this particular dictionary and other than the first priority after-translation expression as result of an after-translation expression is selected from this particular dictionary in a course of the last performed execution of the after-translation expression selection process. This rule implies that, if described by using the case associated with the above mentioned 'facility', the priority levels of User dictionary 1 and User dictionary 2 must be changed when and if the after-translation expression selection process is completed and the situation indicated in drawing (c) shown in Fig.7 is reached with respect to the associated frequency count table.



[0055] In particular, the frequency count connecting to User dictionary 1 and the first priority after-translation expression is '17', and this frequency count is greater than the comparable frequency count '8' connecting to User dictionary 2, which is rated immediately above User dictionary 1 in terms of the priority ranking. This is a situation in which condition-1 is met. Apart from this, the number of times in which after-translation expressions have been selected from User dictionary 1 in the course of the after-translation expression selection process being executed is '3' and this number is greater than the frequency count '2', which is the frequency count connecting to this dictionary and other than the first priority after-translation expression. This is a situation in which condition-2 is met. As both condition-1 and condition-2 having been met in this way, the priority ranking has to be changed so as to assign a higher priority to User dictionary 1 than a priority assigned to User dictionary 2.

[0056] Incidentally, condition-1 and condition-2 mentioned here as conditions used for determining whether any change is required in the priority ranking are merely an example set of conditions and it is of course possible to employ various other conditions for this objective. For example, while this represents a more complicated set of conditions, it is possible to configure an apparatus so as to determine the priority ranking on a basis of rates of respective results of executing a user selection process versus the total number of performed runs of this user selection process wherein the user selection process is one for handling a situation in which a variety of after-translation expressions are found for one common headword in every two pairs of dictionaries.

[0057] It is possible to adopt a rather simple set of conditions for the priority determination, in which the priority is determined on a basis of changes caused by execution of the after-translation expression selection process in the frequency counts connecting to respective dictionaries and the first priority after-translation expressions. And according to this configuration, the priority ranking evaluation and revising process is terminated when it is determined not to require a change as result of performing the determination in Step S502. But when it is determined to require a change in the priority ranking, it moves to Step S503, at which the priority setting unit 12 performs a priority changing process and terminates the priority evaluation and revision process. This implies, in association with the above mentioned case, that the priority order (going down in this order) 'User

dictionary 1, User dictionary 2' is changed to the priority order 'User dictionary 2, User dictionary 1'.

[0058] Above described are all concerned with an objective of reducing work loads associated with the backend processing of work of translating sentences and are directed to a configuration for adjusting the priority ranking of dictionaries automatically while a user progresses with an interactive part of translation processes.

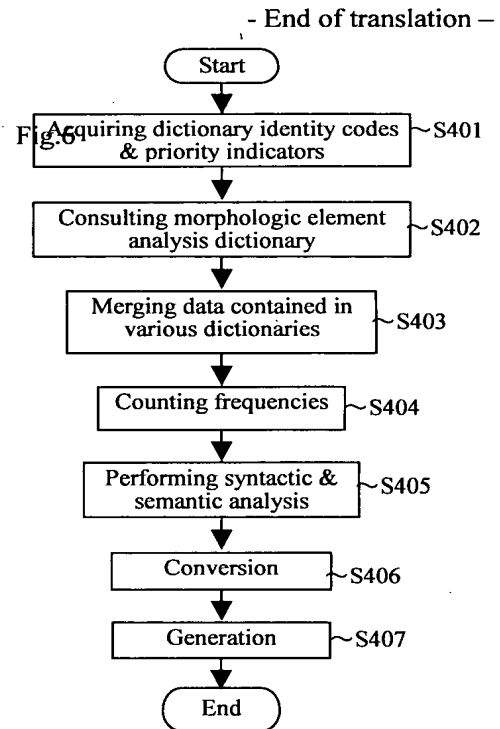
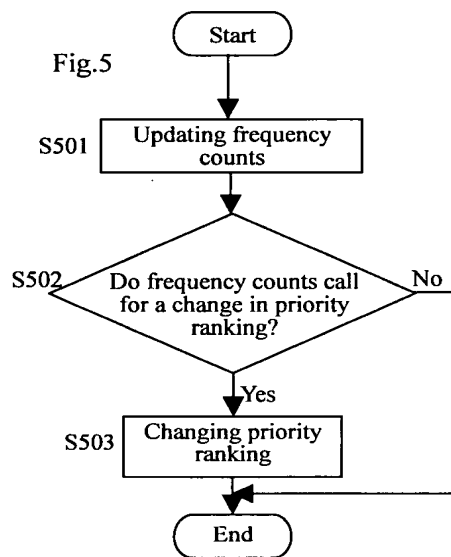


Fig.7

(a)

Dictionary identities	First priority after-translation expression	Other than the first priority after-translation expression	Number of times having been selected in an after-translation expression selection run
Specialty sector dic. 1	21	0	0
Specialty sector dic. 2	12	9	0
Specialty sector dic. 3	5	11	0
User dictionary 2	11	0	0
User dictionary 1	14	5	0

(b)

Dictionary identities	First priority after-translation expression	Other than the first priority after-translation expression	Number of times having been selected in an after-translation expression selection run
Specialty sector dic. 1	21	0	0
Specialty sector dic. 2	12	9	0
Specialty sector dic. 3	5	11	0
User dictionary 2	10	1	0
User dictionary 1	15	4	1

(c)

Dictionary identities	First priority after-translation expression	Other than the first priority after-translation expression	Number of times having been selected in an after-translation expression selection run
User dictionary 2	8	3	0
User dictionary 1	17	2	3

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-81572

(43)公開日 平成9年(1997)3月28日

(51)Int.Cl.<sup>6</sup>

G 0 6 F 17/28

識別記号

庁内整理番号

F I

G 0 6 F 15/38

技術表示箇所

C

T

審査請求 未請求 請求項の数 2 O L (全 11 頁)

(21)出願番号 特願平7-234030

(22)出願日 平成7年(1995)9月12日

(71)出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72)発明者 吉村 裕美子

東京都府中市東芝町1番地 株式会社東芝

府中工場内

(72)発明者 木下 聡

神奈川県川崎市幸区小向東芝町1番地 株

式会社東芝研究開発センター内

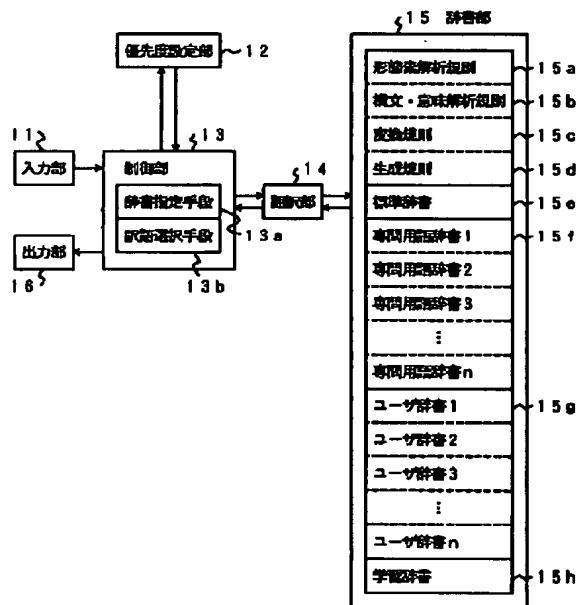
(74)代理人 弁理士 鈴江 武彦

(54)【発明の名称】 翻訳装置及び辞書優先度設定方法

(57)【要約】

【課題】ユーザの訳語選択の結果を学習材料として、ユーザの意識しないうちに、初期の優先度指定内容をより翻訳対象文書に適応した優先度に調整し、訳語選択・後編集の手間を軽減する。

【解決手段】辞書部15に予め用意された各辞書の中でオプショナル辞書として指定された辞書の種類とその優先度に従って、翻訳部14により翻訳対象となる第1の言語が第2の言語に翻訳処理される。その際に、優先度設定部12によって翻訳結果として得られる訳語の選択内容に基づいて優先度の変更の必要性が判断され、優先度の変更が必要である場合に、最も訳語選択の多い辞書を優先するように、その優先度の調整が行われる。これにより、ユーザの訳語選択により、常に最適な優先度で複数辞書の登録訳語を訳文に生成することができ、その結果、ユーザの訳語選択・後編集の手間を軽減することができる。



## 【特許請求の範囲】

【請求項1】 自然言語を他の言語に変換するための複数の辞書を記憶した辞書記憶手段と、

翻訳対象となる第1の言語を入力する入力手段と、

上記辞書記憶手段に記憶された各辞書を用いて、上記入力手段によって入力された第1の言語を第2の言語に翻訳処理する翻訳手段と、

この翻訳手段によって得られた翻訳結果を出力する出力手段と、

上記辞書記憶手段に記憶された各辞書の中でオブショナル辞書の種類とその優先度を指定する辞書指定手段と、  
上記出力手段によって出力された翻訳結果の中から訳語を選択する訳語選択手段と、

上記辞書指定手段によって指定されたオブショナル辞書の種類とその優先度に基づいて上記翻訳手段が翻訳処理で用いる辞書の優先度を決定し、かつ、上記訳語選択手段によって選択された訳語の選択内容に基づいて優先度の変更の必要性を判断し、優先度の変更が必要である場合に、オブショナル辞書の優先度の調整を行なう優先度設定手段とを具備したことを特徴とする翻訳装置。

【請求項2】 自然言語を他の言語に変換するための複数の辞書を記憶した辞書メモリを有し、

この辞書メモリに記憶された各辞書の中でオブショナル辞書として指定された辞書の種類とその優先度に従って、翻訳対象となる第1の言語を第2の言語に翻訳処理し、

その翻訳結果として得られる訳語の選択内容に基づいて優先度の変更の必要性を判断した後、

優先度の変更が必要である場合に、オブショナル辞書の優先度の調整を行なうにしたことを特徴とする辞書優先度設定方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】この発明は、第1言語の文を第2言語の文に翻訳するための翻訳装置に係り、特に標準辞書の他に、ユーザ辞書、専門用語辞書をオブショナル辞書として翻訳処理に用いる際に好適な翻訳装置及び辞書優先度設定方法に関する。

## 【0002】

【従来の技術】近年、コンピュータを利用して、ある言語の文章を自動的に別の言語の文章に変換する機械翻訳システムに対する注目が高まっている。現在利用されている機械翻訳システムは、システム固有の標準辞書の他に、ユーザ辞書、専門用語辞書をオブショナル辞書として利用できる環境を提供しているのが通常である。また、ユーザ辞書、専門用語辞書のそれぞれについても複数個を使った翻訳を可能にしているシステムも多い。

【0003】このような各種の辞書を用いて翻訳処理を行う場合には、基本的には、翻訳対象文書の内容に従ってユーザが辞書間の優先度を指定できる。これにより、

ある見出し語が複数の辞書にまたがって登録されている場合には、翻訳システム側はユーザの指定に従い、優先度の高い辞書に登録されている訳語を訳文に生成するようにする。

【0004】ところが、常にユーザが翻訳対象の個々文書の内容を把握しきれているわけでもなく、また、文書の内容が異なる毎に辞書の優先度を指定をし直すこともユーザには手間となる。このため、結果として、内容的に優先度を低くすべき辞書の訳語が訳出されてしまうことは回避しきれない。

【0005】これに対しては、ユーザは「訳語選択」の手段を行ない、訳語の学習を行なうことで対処していた。「訳語選択」とは、翻訳結果として出力された複数の訳語候補の中で所望の訳語を選択することである。この選択された訳語を記憶（学習）しておくことで、以後、同じ単語が入力された際に、その訳語を優先出力することがきる。

## 【0006】

【発明が解決しようとする課題】上記したように、従来、標準辞書の他に、ユーザ辞書、専門用語辞書といったオブショナル辞書をそれぞれ複数有して翻訳処理を行う場合には、その都度、ユーザが各辞書の優先度を指定していた。

【0007】しかしながら、常に所望の訳語を得るには、翻訳対象となる文書の内容を把握し、その文書内容毎に優先度を指定する必要があるため、かなりの困難を要することになる。

【0008】また、「訳語選択」によって対処したとしても、根本的にオブショナル辞書の最初の優先度指定が不適切であると、何度となく「訳語選択」を行なうことになる。このため、ユーザに負担をかける等の問題があった。

【0009】本発明は上記のような点に鑑みなされたもので、ユーザの訳語選択の結果を学習材料として、ユーザの意識しないうちに、初期の優先度指定内容をより翻訳対象文書に適応した優先度に調整し、訳語選択・後編集の手間を軽減することのできる翻訳装置及び辞書優先度設定方法を提供することを目的とする。

## 【0010】

【課題を解決するための手段】上記目的を達成するため、本発明は、自然言語を他の言語に変換するための複数の辞書を記憶した辞書記憶手段と、翻訳対象となる第1の言語を入力する入力手段と、上記辞書記憶手段に記憶された各辞書を用いて、上記入力手段によって入力された第1の言語を第2の言語に翻訳処理する翻訳手段と、この翻訳手段によって得られた翻訳結果を出力する出力手段と、上記辞書記憶手段に記憶された各辞書の中でオブショナル辞書の種類とその優先度を指定する辞書指定手段と、上記出力手段によって出力された翻訳結果の中から訳語を選択する訳語選択手段と、上記辞書指定

手段によって指定されたオブショナル辞書の種類とその優先度に基づいて上記翻訳手段が翻訳処理で用いる辞書の優先度を決定し、かつ、上記訳語選択手段によって選択された訳語の選択内容に基づいて優先度の変更の必要性を判断し、優先度の変更が必要である場合に、オブショナル辞書の優先度の調整を行なう優先度設定手段とを具備したことを特徴とする。

【0011】このような構成によれば、予め用意された各辞書の中でオブショナル辞書として指定された辞書の種類とその優先度に従って、翻訳対象となる第1の言語が第2の言語に翻訳処理される。

【0012】その際に、翻訳結果として得られる訳語の選択内容に基づいて優先度の変更の必要性が判断され、優先度の変更が必要である場合に、最も訳語選択の多い辞書を優先するように、その優先度の調整が行われる。

【0013】これにより、ユーザの訳語選択により、常に最適な優先度で複数辞書の登録訳語を訳文に生成することができ、その結果、ユーザの訳語選択・後編集の手間を軽減することができる。

【0014】

【発明の実施の形態】以下、図面を参照して本発明の一実施形態を説明する。図1は本発明の一実施形態に係る翻訳装置の構成を示すブロック図である。本装置は、入力部11、優先度設定部12、制御部13、翻訳部14、辞書部15、出力部16を有する。

【0015】入力部11は制御部13に接続され、制御部13は出力部16に接続される。また、制御部13は、辞書指定手段13a、訳語選択手段13bを有し、優先度設定部12および翻訳部14と双方向に接続されている。

【0016】さらに、翻訳部14は辞書部15と双方向に接続されている。辞書部15は、形態素解析規則15a、構文・意味解析規則15b、変換規則15c、生成規則15d、標準辞書15e、専門用語辞書15f、ユーザ辞書15g、学習辞書15hからなる。学習辞書15hとは、ユーザの訳語選択により学習された訳語の情報を収納する辞書である。

【0017】入力部11は、制御部13に対して、翻訳対象となる文の入力や各種コマンドの入力を行なうためのものである。通常、この入力部11としては、キーボード、マウスなどが用いられる。また、翻訳対象文の入力には、OCRの他、フロッピーディスク、磁気テープ、磁気ディスクからの読み込みも考えられる。

【0018】出力部16は、翻訳部14によって得られた翻訳結果を出力したり、各種命令に対する制御部13の応答を表示するためのものである。この出力部16としては、CRT (Cathode Ray Tube) やLCD (Liquid Crystal Display) 等の各種ディスプレイ装置が通常であるが、翻訳結果の出力には、印刷機などの印字手段、あるいは直接フロッピーディスク、磁気テープ、磁気デ

ィスクへ出力する手段や、他のメディアへ送信する送信手段でも良い。

【0019】制御部13は、本装置全体の制御を行うものである。辞書指定手段13aは、翻訳処理に用いるオブショナル辞書とその種類を指定するためのものである。訳語選択手段13bは、翻訳処理によって得られた翻訳結果の中から訳語を選択するためのものである。

【0020】このような構成において、ユーザは、入力部11より制御部13の辞書指定手段13aを介して、翻訳処理に用いるオブショナルの専門用語辞書15f、ユーザ辞書15gの種類を指定する命令を制御部13に送る。専門用語辞書15f、ユーザ辞書15gをそれぞれ複数個指定する際には、その辞書間の優先度の指定も制御部13に送る。

【0021】ユーザから辞書の指定がなされると、その指定内容は制御部13を通じて優先度設定部12に送られる。これにより、優先度設定部12は、初期設定としてユーザの指定内容を保持する。

【0022】入力部11により翻訳対象の原文が入力されると、その入力データは制御部13に送られる。制御部13は、原文のデータを翻訳部14に送る。翻訳部14は、制御部13を通じて優先度設定部12からオブショナル辞書の種類とその優先度の情報を得る。そして、翻訳部14は、その情報と辞書部15の知識・規則を利用して、入力原文に対して第2言語の文あるいは第1言語の文字列を含んだ第2言語の文への翻訳処理を行なう。

【0023】翻訳処理が終了すると、その翻訳結果は制御部13を通じて出力部16に送られ、一通りの処理を終える。ここで、翻訳結果が得られた状態で、入力部11より制御部13の訳語選択手段13bを介してある語に対する訳語選択の命令が制御部13に送られると、制御部13は、ユーザの選択内容に従った表示命令を出力部16に送る。それと同時に、訳語選択の内容が優先度設定部12にも送られる。

【0024】この訳語選択により、どの辞書の訳語が選択・学習され、どの辞書の訳語が置き換えられたか、現状でどの辞書の訳語が訳文に生成されているかなどをもとに、優先度設定部12は現状の辞書間の優先度の再評価を行なって、最適な優先度を再設定する。

【0025】それ以降の翻訳処理は、訳語選択が行なわれる度に、新規の優先度に従って訳語設定が行なわれる。次に、制御部13における対話的な翻訳処理の流れについて説明する。

【0026】図2はその対話的な翻訳処理の流れを示すフローチャートである。翻訳に際し、まず、ユーザによりオブショナル辞書として使用する辞書の指定が行なわれる(ステップS201)。

【0027】指定の内容は優先度設定部12に送られ、辞書の種類、辞書間の優先度が初期値として保持される

10

20

30

40

50

(ステップS202)。この段階で、優先度設定部12は、後述するステップS215の辞書間優先度更新の処理に用いる頻度カウントテーブルを初期化する。この頻度カウントテーブルの扱いについては、後に詳しく説明する。

【0028】以降、制御部13は、入力部11から何等のキー入力があると(ステップS203)、ステップS204~209にて、キー・コマンドに応じた処理を実行する。

【0029】すなわち、例えば入力部11より終了命令が入力されると(ステップS204のYes)、制御部13は全ての処理を終了する。また、入力部11より翻訳指示命令を受けた場合には(ステップS205のYes)、ステップS210へ進み、制御部13は指示対象文に対する翻訳処理開始命令を翻訳部14に送る。このときの翻訳部14における翻訳処理の流れについては図4を参照して後に説明する。

【0030】翻訳処理が終了するとステップS211へ進み、制御部13は出力部16に対して訳文を表示する命令を送る。また、入力部11より文字キーの入力があった場合には(ステップS206のYes)、ステップS212へ進み、制御部13は原文の文字列として出力部16に表示命令を送る。

【0031】また、入力部11より編集キーの入力があった場合には(ステップS207のYes)、ステップS213へ進み、制御部13は編集処理を行なう。入力部11より訳語選択の命令が入力された場合には(ステップS208のYes)、ステップS214へ進み、制御部13はユーザに訳語選択を促す処理を行なう。

【0032】図3は出力部16に表示してユーザの訳語選択処理をガイドする画面の一例である。この例は、英日翻訳において英単語「facility」に対する訳語の一覧を表示し、ユーザの選択を委ねる画面である。

【0033】訳語の右に表示される「T1」、「T2」、「T3」はそれぞれ専門用語辞書1、2、3に登録されている訳語であることを示す。同様に、「U1」、「U2」はユーザ辞書1、2に登録されていることを示す。残る「C」は標準辞書に登録されていることを示す。

【0034】しかして、ステップS214において、訳語の選択が行なわれると、制御部13は選択内容を学習辞書15hに保存すると共に、ステップS215へ進み、その選択内容を優先度設定部12に送る。

【0035】優先度設定部12では、その内容に基づいて辞書間の優先度の再評価を行ない、優先度の更新を行なう。この詳細については、後に図5を参照して詳しく説明する。

【0036】次に、翻訳部14における翻訳処理の流れを説明する。図4はその翻訳処理の流れを示すフローチャートである。まず、翻訳部14は、制御部13を経て

優先度設定部12が保存しているオプショナル辞書の種類情報と辞書間の優先度情報を入手する(ステップS401)。

【0037】これらの情報が入手されると、翻訳部14は、辞書部15に記憶された形態素解析規則15a、標準辞書15e、さらに優先度設定部12より得たオプショナル辞書の種類情報に基づいて専門用語辞書15f、ユーザ辞書15gを選出し、形態素解析・辞書引き処理を行なう(ステップS402)。

【0038】この段階では、入力文中に複数の辞書に登録されている語句があると、それぞれの辞書からの登録情報が検索される。そこで、続くステップS403では、複数の辞書から得られた登録内容をマージする処理を行なう。

【0039】すなわち、翻訳部14は、優先度設定部12から得た辞書間の優先度情報に基づいて、一つの見出し語に対する複数の登録内容を、あたかも1つの辞書から得られた登録内容であるかのようにマージを行なう。

【0040】以下に、オプショナル辞書の種類・優先度とそれに応じた見出し語「facility」の登録内容のマージ結果の例を挙げる。\*オプショナル辞書の種類

専門用語辞書1、2、3

ユーザ辞書1、2

\*辞書間優先度(優先度順に列挙)

専門用語辞書1、2、3

ユーザ辞書2、1

\*登録内容

専門用語辞書1

訳語=機能、施設、装置

専門用語辞書2

訳語=機能

専門用語辞書3

訳語=装置、機構

ユーザ辞書1

訳語=ファシリティ

ユーザ辞書2

訳語=設備

標準辞書

訳語=機能、施設、機構、設備、能力

\*マージ結果

訳語=設備(U2、C)

ファシリティ(U1)

機能(T1、T2、C)

施設(T1、C)

装置(T1、T3)

機構(T3、C)

能力(C)

なお、T1~T3は専門用語辞書、U1、U2はユーザ辞書、Cは標準辞書を示す。

【0041】上記例では、一般にユーザ辞書の登録内容は専門用語辞書より優先度を高めるという仕様をとっており、辞書間優先度については専門用語辞書はその専門用語辞書同志、ユーザ辞書はそのユーザ辞書同志のみ優先度を指定するようにしている。当然ながら、専門用語辞書とユーザ辞書とを区別せずに、一元化して優先度を指定できるようにする実施方法も可能である。

【0042】ここで、ユーザにより別に行なわれた訳語選択による訳語の学習情報が学習辞書15hにある場合は、翻訳部14は上記マージ後に第一訳語（ここでは「設備（U2, C）」が学習された訳語と同じであるかのチェックを行ない、第一訳語以外の訳語が学習された訳語と一致する場合には、訳語のマージ結果に対して、訳語の順番の入れ替えを行なう。

【0043】また、学習された訳語がマージ結果の中に入らない場合は、翻訳部14は学習された訳語をマージ結果の最初に挿入する処理を行なう。例えば、学習された訳語が「機能」である場合には、以下ようになる。

【0044】

＊マージ結果

訳語＝機能（T1, U2, C）  
設備（U2, C）  
ファシリティ（U1）  
施設（T1, C）  
装置（T1, T3）  
機構（T3, C）  
能力（C）

また、学習された訳語が「便宜」である場合には、以下のようなになる。

【0045】

＊マージ結果

訳語＝便宜、設備（U2, C）  
ファシリティ（U1）  
機能（T1, T2, C）  
施設（T1, C）  
装置（T1, T3）  
機構（T3, C）  
能力（C）

このようにして、辞書情報のマージが終わると、翻訳部14は辞書毎に原文中に登録語が現われるカウントを保存した頻度カウントテーブルをアップデートする（ステップS404）。

【0046】なお、この頻度カウントテーブルは、翻訳部14の中の図示せぬ作業領域に設けられており、ここでは翻訳部14を通じてアップデートされる。また、初期時において、この頻度カウントテーブルは、図2のステップS202で初期化されている。

【0047】図7（a）は頻度カウントテーブルのこの時点での内容を示す図である。ここに示されているように、第一訳語とその他の訳語とで区別してカウントす

る。ただし、第一訳語にも第二訳語以降にも同じ辞書の訳語が登録されている場合には「その他の訳語」としては計上しない。このような計上を翻訳終了の命令が指示されるまで累計して行なう。

【0048】その後、翻訳部14は構文・意味解析規則15bを用いて入力文の構文・意味解析を行ない（ステップS405）、続いて、変換規則15cを用いて解析結果から第2言語の構造への変換を行なう（ステップS406）。そして、最後に、翻訳部14は生成規則15dを用いて第2言語の文章として生成を行ない、一単位の翻訳処理を終了する（ステップS407）。

【0049】次に、訳語選択後に行なわれる優先度設定部12における辞書間優先度再評価・更新の処理の流れを説明する。図5はその処理の流れを示すフローチャートである。ユーザにより第一訳語以外の訳語が訳語選択されると、優先度設定部12はその新たに選択された訳語が登録されていた辞書の種類と、もとの訳語が登録されていた辞書の種類の情報に基づいて、辞書毎の上記頻度カウントテーブルの内容をアップデートする（ステップS501）。

【0050】上記「facility」の例を用いれば、ユーザ辞書2の訳語が第一訳語であり、それが訳文にも生成されている状態で、訳語選択によりユーザ辞書1に登録されている「ファシリティ」が選択されたとすると、頻度カウントテーブルは図7（b）に示すように書き換えられる。

【0051】この場合、ユーザ辞書1の第一訳語のカウント数を「14」から「15」にカウントアップし、その他の訳語のカウント数を「5」から「4」にカウントダウンする。これに対し、ユーザ辞書2の第一訳語のカウント数を「11」から「10」にカウントダウンし、その他の訳語のカウント数を「0」から「1」にカウントアップする。

【0052】このようにして、頻度カウントテーブルがアップデートされると、優先度設定部12はその内容をもとに辞書間の優先度の変更が必要かどうかを判定する（ステップS502）。

【0053】本実施形態では、以下の条件を全て満たすときに優先度の変更を行なうものとする。条件1. 直前の訳語選択において選択された訳語が登録されている辞書の第一訳語の総数が、優先度がひとつ上位の辞書より多い。

【0054】条件2. 直前の訳語選択において選択された訳語が登録されている辞書の訳語選択による選択回数がある他の訳語の数より多い。上記「facility」の例を用いれば、訳語選択により頻度のカウントテーブルが図7（c）に示すような状況になったとき、ユーザ辞書1とユーザ辞書2の間の優先度の変更が必要であると判断する。

【0055】すなわち、ユーザ辞書1の第一訳語のカウ



ント数は「17」であり、優先度がひとつ上のユーザ辞書2のカウンタ数である「8」より大きいため、条件1を満たしている。また、ユーザ辞書1に関しては、訳語選択により訳語が選択された回数が3回であり、訳語が第1訳語以外のものが選択された回数である2回より大きいため、条件2を満たしている。これにより、2つの条件が満たされているため、ユーザ辞書1をユーザ辞書2より優先させる処理を行う。

【0056】なお、ここで挙げた優先度変更の必要性判断に用いた条件はあくまでも一例であり、他の基準を設定することも当然ながら可能である。例えば、より複雑には、2対の辞書毎に同じ見出し語で異なる訳語が登録されている際の訳語で、対立全てに対してユーザによる選択結果の比率を基準にして優先度を判定することもできる。

【0057】また、より簡易には、訳語選択による第一訳語の総数の変化のみをキーにして優先度を判定することもできる。ステップS502での判定の結果、優先度の変更の必要はないと判定された場合はそのまま処理を終了する。優先度の変更の必要があると判定された場合は、ステップS503へ進み、優先度設定部12は優先度の変更処理を行ない、優先度評価・更新の処理を終了する。具体的には、上記例でユーザ辞書間の優先度を「ユーザ辞書2、ユーザ辞書1」の順序であったのを、「ユーザ辞書1、ユーザ辞書2」へと変更する。

【0058】以上は、対話的に翻訳処理を行なっていく上で辞書間の優先度を自動調節することにより、以降の文の翻訳における後処理作業の手間を軽減することについて説明してきた。

【0059】このような対話的翻訳処理に対比する翻訳作業の進め方として、最初に一括的に翻訳を行ない、その後で、対話的に後編集を行なうという方法がある。後編集は通常文書の先頭から後方へと進めていくことが多い。

【0060】そこで、後編集時の訳語選択操作については、辞書間優先度判定のキーになる第一訳語、その他の訳語のカウンタは訳語選択の対象語のうち最も文書の後方にある語の位置から文書の先頭位置までの範囲を対象とする。

【0061】ある語に対する訳語選択と共に辞書間優先度の変更の必要性があると判定され、優先度の更新がされた場合は、最後方の訳語選択対象語以降の訳文において、優先度の変更に応じて訳語のチェックを行ない、必要ならば訳語の変更を行なう。

【0062】以下、図6を参照して後編集時での処理について説明する。図6は後編集時に訳語選択が行なわれた後の制御部13の流れを示すフローチャートである。ある語に対して訳語選択が行われると、制御部13は、まず、その語の位置が「最後方」であるか否かをチェックする(ステップS601)。

【0063】「最後方」とは、訳語選択を行った語を含む文の中で文頭から最も遠い文を指す。したがって、ステップS601で、「訳語選択しようとしている語の位置が最後方である」とは、その後を含む文が、その時点での最後方の文より後方に位置し、訳語選択後、新たな「最後方」文になることを示す。

【0064】なお、後編集を開始して初めて訳語選択を行う場合には、「最後方」文は存在しないため、便宜上文書の先頭を仮想的に最後方文とみなす。ここで、後編集における頻度カウンタテーブルは、文書の先頭から最後方の文までの範囲で作成するようにする。

【0065】「訳語選択しようとしている語の位置が最後方である」場合には、制御部13は訳語選択対象文の位置を訳語選択・優先度処理変更後に新たな最後方文として設定するため、その位置を記憶しておく。さらに、制御部13は、まだ頻度カウンタテーブルに反映されていない部分、すなわち、その時点での最後方の文の次の文から、その語を含む文までの範囲を参照し、頻度カウンタテーブルをアップデートする(ステップS602)。

【0066】その後、制御部13は、実際に行われた訳語の選択の結果に基づいて、頻度カウンタテーブルのアップデートを行う(ステップS603)。次に、制御部13は、当該語の位置にかかわらず、頻度カウンタテーブルの内容に基づいて、辞書間優先度の変更が必要かどうかを判定する(ステップS604)。

【0067】必要であると判定された場合、制御部13は優先度設定部12を通じて優先度の変更を行ない(ステップS605)、続いて、訳語選択対象語のうち最後方の語の位置より後方の訳文に関して、オプション辞書の訳語が生成されている箇所を検索し、その訳語が変更された優先度に適合するものかどうかをチェックする(ステップS606)。適合しない場合には、優先度にそった訳語と入れ替えを行なう。

【0068】最後に、制御部13は訳語選択対象語のうちの最後方の語の位置より後方に訳文に対して、当該訳語選択による訳語の学習の内容の反映を行なって全ての処理を終了する(ステップS607)。

【0069】なお、上記実施形態では、後編集におけるユーザの着眼点(後処理の対象語)の移動は、訳語選択によるもの以外は対象から外しているが、他の編集作業の対象も含めて着眼点の最後方語を定めてもよい。

【0070】また、上記実施形態では、辞書間優先度の変更や訳語選択による訳語の学習があると、着眼点の最後方の位置より後の訳文に対して(必要ならば)訳語の入れ替えを自動的に行なうようにしているが、自動的には行わずに、再翻訳の指示があった文に対してのみ変更された優先度の内容、訳語学習の内容を反映するようにする実施の方法も可能である。

【0071】以上説明したように、オプション辞書間

の優先度の初期設定が実際の文書の内容とは適合していない場合でも、ユーザが訳語選択をすることにより、意識せずとも文書に合った優先度へと自動調整することができる。この場合、訳語選択による訳語学習の内容は保存されているので、優先度の自動調整が選択した訳語に対して副作用を及ぼすこともない。

【0072】

【発明の効果】以上のように本発明によれば、複数のオプション辞書を使用して翻訳処理を行う場合において、指定されたオプション辞書の種類とその優先度に基づいて翻訳処理で用いる辞書の優先度を決定後、訳語選択の内容に基づいて優先度の変更の必要性を判断し、優先度の変更が必要である場合に、最も訳語選択の多い辞書を優先するように、その優先度の調整を行なうようにしたため、オプション辞書間の優先度の初期設定が実際の文書の内容とは適合していない場合であっても、ユーザが訳語選択を繰り返すうちに当該文書に適した優先度を自動的に導出できる。これにより、訳語選択毎に優先度が調整されていくにつれて、初期設定の不適切さに起因する訳語選択の手間を軽減することができる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係る機械翻訳装置の構成を示すブロック図。

【図2】上記機械翻訳装置の制御部における対話的翻訳処理の流れを示すフローチャート。

【図3】上記機械翻訳装置においてユーザの訳語選択処

\* 理をガイドする画面の一例を示す図。

【図4】上記機械翻訳装置の翻訳部における翻訳処理の流れを示すフローチャート。

【図5】上記機械翻訳装置の優先度設定部における処理の流れを示すフローチャート。

【図6】上記機械翻訳装置の制御部における後編集時の処理の流れを示すフローチャート。

【図7】上記機械翻訳装置で用いられる頻度カウンテーブルの内容を示す図。

【符号の説明】

- 11…入力部、
- 12…優先度設定部、
- 13…制御部、
- 13a…辞書指定手段、
- 13b…訳語選択手段、
- 14…翻訳部、
- 15…辞書部、
- 15a…形態素解析規則、
- 15b…構文・意味解析規則、
- 15c…変換規則、
- 15d…生成規則、
- 15e…標準辞書、
- 15f…専門用語辞書、
- 15g…ユーザ辞書、
- 15h…学習辞書、
- 16…出力部。

【図3】

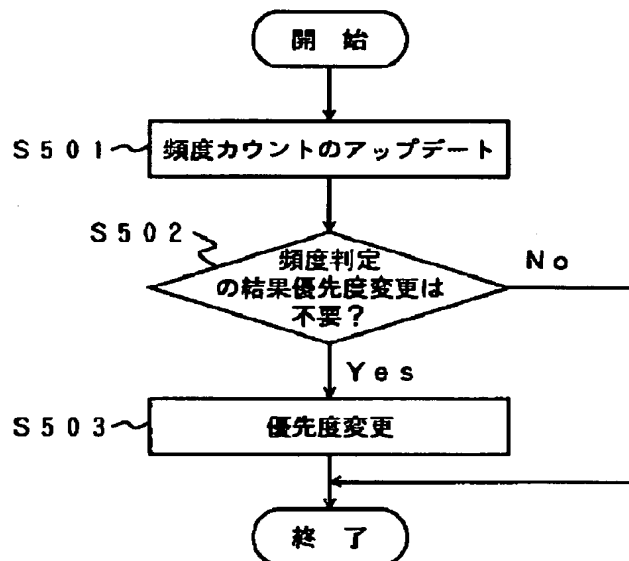
<訳語選択>
前ページ
次ページ

希望の訳語をクリックして下さい。

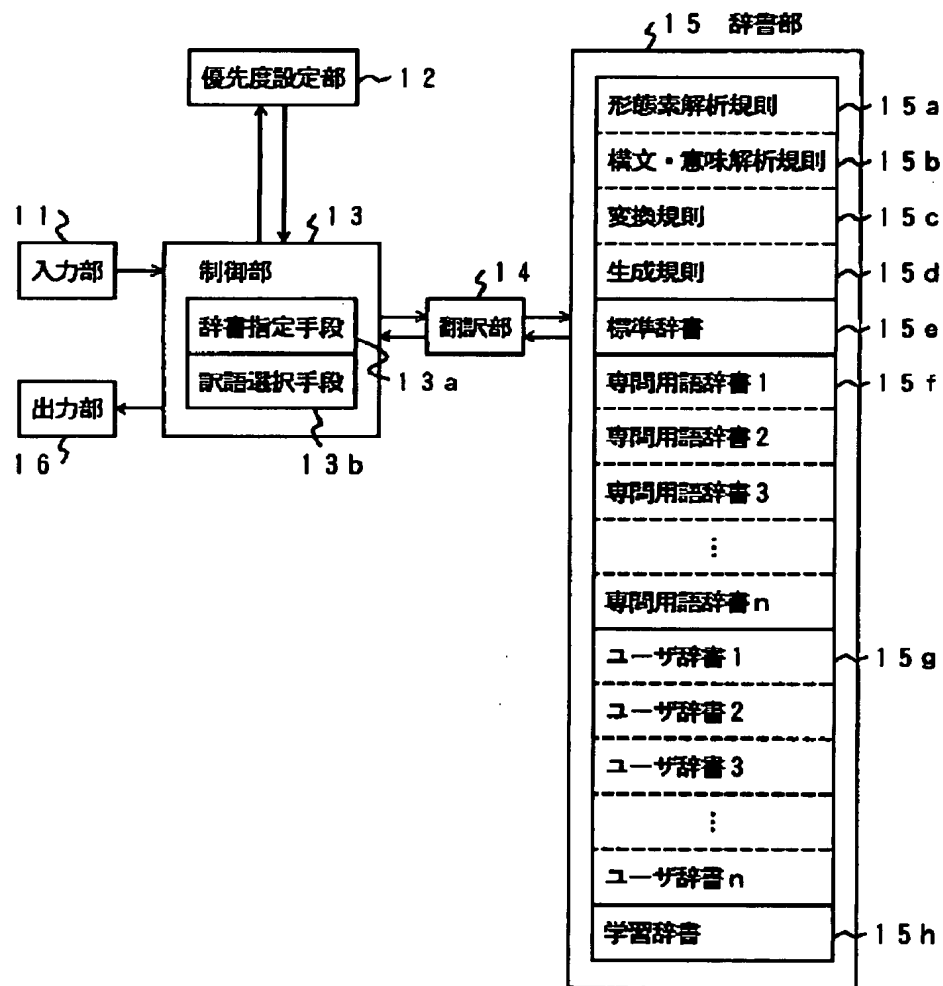
見出し語「facility」

<input type="checkbox"/> 機能	(T1, T2, C)
<input type="checkbox"/> 施設	(T1, C)
<input type="checkbox"/> 装置	(T1, T3)
<input type="checkbox"/> 機構	(T3, C)
<input type="checkbox"/> ファシリティ	(U1)
<input type="checkbox"/> 設備	(U2, C)
<input type="checkbox"/> 能力	(C)

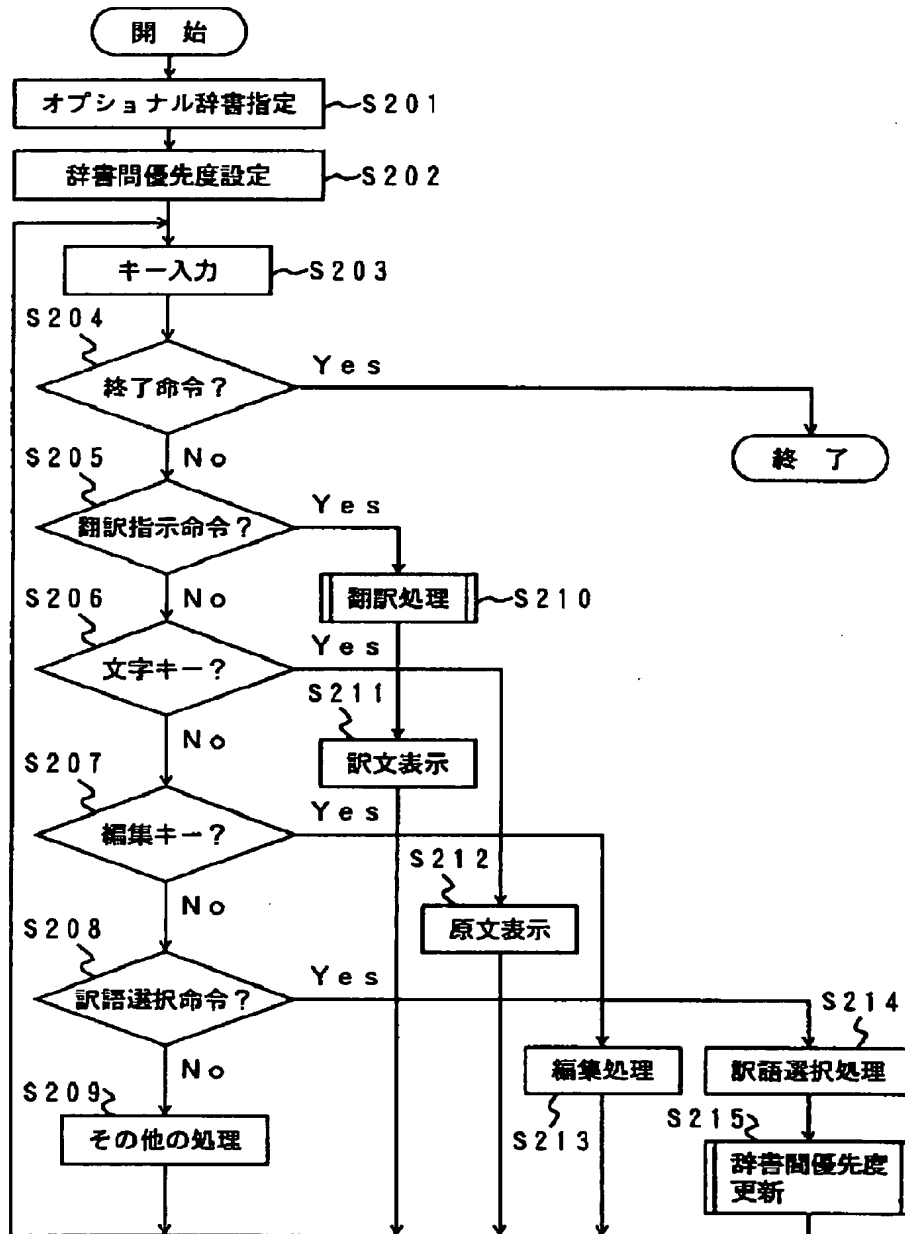
【図5】



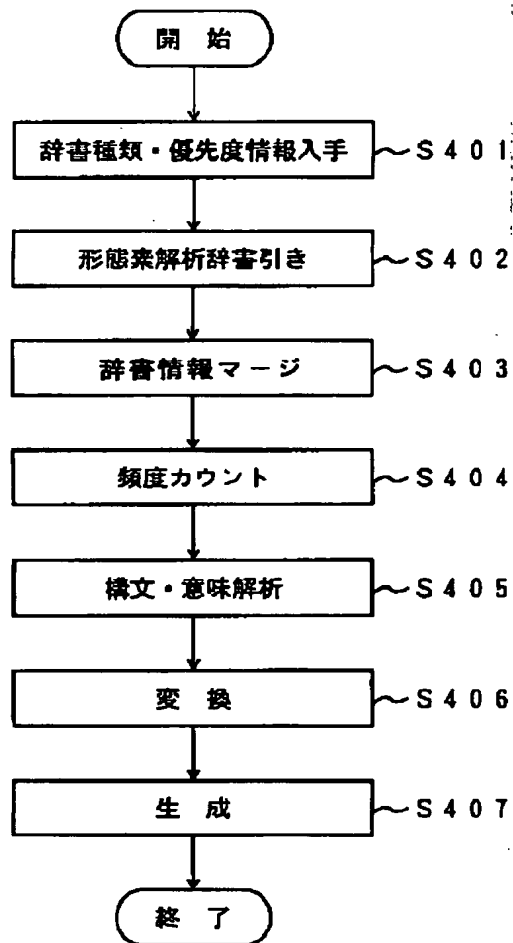
【図1】



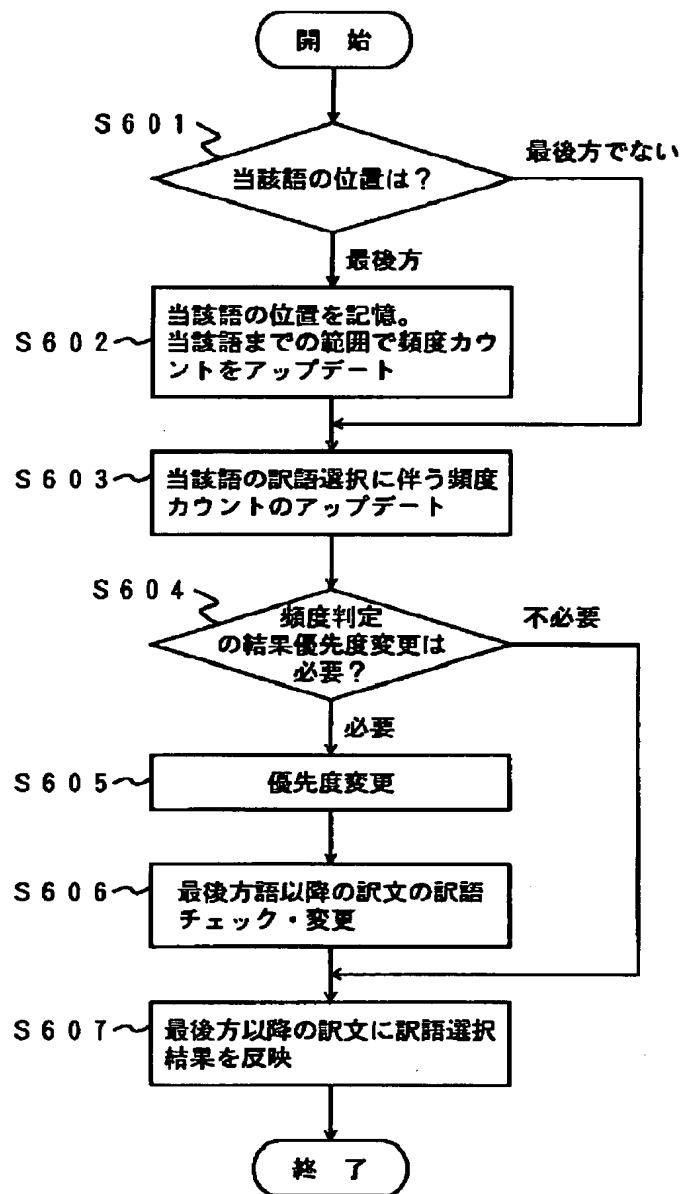
【図2】



【図4】



【図6】



【図7】

(a)

辞書種類	第一訳語	その他の訳語	訳語選択による選択回数
専門用語辞書1	21	0	0
専門用語辞書2	12	9	0
専門用語辞書3	5	11	0
ユーザ辞書2	11	0	0
ユーザ辞書1	14	5	0

(b)

辞書種類	第一訳語	その他の訳語	訳語選択による選択回数
専門用語辞書1	21	0	0
専門用語辞書2	12	9	0
専門用語辞書3	5	11	0
ユーザ辞書2	10	1	0
ユーザ辞書1	15	4	1

(c)

辞書種類	第一訳語	その他の訳語	訳語選択による選択回数
ユーザ辞書2	8	3	0
ユーザ辞書1	17	2	3